ISSN: 2998-2766
Iris Online Journal of Sciences
DOI: 10.33552/IOJS.2024.01.000511
Iris Publishers

Research Article

Selection of Potential Microsatellite Markers for Marker-Trait Associations in Tropical Sweet Corn

Pedram Kashiani*

Faculty of Science, Universiti Tunku Abdul Rahman (UTAR), Kampar 31900 Kampar, Perak, Malaysia

***Corresponding author:** Pedram Kashiani, Faculty of Science, Universiti Tunku Abdul Rahman (UTAR), Kampar 31900 Kampar, Perak, Malaysia.

Received Date: July 02, 2024

Published Date: July 26, 2024

Abstract

Most of the agronomic traits in sweet corn (Zea mays cover. saccharata) are controlled by several genes at different loci with varying degrees of genetic contribution. In the current study, a preliminary identification of genomic regions responsible for agronomic traits in tropical sweet corn was conducted. For this purpose, a genome-wide association study was conducted on 13 tropical inbred lines using 150 microsatellite DNA markers. A total of 1347 alleles were amplified, out of which 344 were unique among the inbred lines. Prior to association mapping, population structure was assessed based on allele frequency using Bayesian model-based approach, identifying five main heterotic groups among the inbred lines. Association analysis was performed with 10,000 permutation replicates to determine chromosome-wise and genome-wise significance levels of association tests. In total, 19 significant marker-trait associations were identified when significance level was set at $p \leq 0.01$. The portion of total variance estimated by the marker ranged from 0.303 to 0.635. Each chromosome possessed at least one (Chromosomes 1, 2 and 6) to three loci (Chromosomes 3 and 7) associated with one of 17 agronomic traits measured. Among the yield related traits, dehusked ear diameter was found to be associated with bnlg1526 located on Chromosome 10, while husked ear length and dehusked ear weight were found to be associated with phi053 (Chr-3) and umc1143 (Chr-6), respectively. These significant associations identified will be utilized as a powerful tool to uncovering potential QTL-allele matrix for further marker-assisted breeding on tropical sweet corn.

Keywords: Marker-trait association; Microsatellite markers; Inbred line; Tropical sweet corn

Introduction

Corn as one of the most phenotypically diverse crop species can be grown in a wide range of environments, so long as its water requirements are in accordance [1]. Complex quantitative traits are usually influenced by a large number of genes as well as environmental effects especially in cross-pollinated crops with higher degree of heterozygosity like corn. Understanding the genetic basis of complex traits have traditionally been the focus of quantitative genetics, which relies on partitioning phenotypic variation within and among individuals with known degrees of relatedness [2]. However, as the availability of useful genetic markers have increased, it has become possible to associate genome regions containing these markers to variation in complex traits. Microsatellites present a relatively new class of DNA-based markers boasting numerous advantages including reliability, high polymorphism, abundance, co-dominance, reproducibility, and somatic stability stable [3-5]. These make them reliable for germplasm characterization, genotyping and fingerprinting in many plant species [6-8].

Genetic linkage map construction has been recognized as an essential tool for plant molecular breeding using DNA markers

because they are neutral, lack epistasis and are simply inherited in Mendelian nature. The method of linkage analysis is well developed for bi-parental crosses between inbred lines. However, one of the limiting factors in genomic analysis of many plant species is that most genomic studies have been conducted in experimental populations developed from a bi-parental cross. Thus, while many quantitative trait loci (QTL) have been reported, the effects of these QTLs often turn out to be unique to a specific genetic background, and there has been limited success in applying the results across breeding populations. Many researchers now consider that association analysis, whereby genes and QTL are detected in a random set of genotypes from a mixed genetic background, is a viable solution to this problem [9]. The increased availability of molecular markers and the refinement of statistical tools have kindled renewed interest in this approach. Association mapping (AM) is based on the assumption that there is a set of markers available and either they represent actual genes (or alleles) or that of the markers are so close to the actual functional genes that they co-segregate and happen to be in linkage disequilibrium (LD). This implies that the LD mapping is done with a natural population in which association between traits and markers exists due to linkage disequilibrium. The degree of LD depends on the recombination events that have taken place in history [10-11].

The goal of this research was to apply AM approaches to identify SSR markers associated with important agronomic traits in 13 advanced sweet corn inbred lines. Pedigree information was not available for the lines. Therefore, there is a potential for false marker-trait associations because of population structure or family relatedness. Statistical procedures that account for population structure and family relatedness were employed to minimize false positives and maximize power.

Materials And Methods

A set of homozygous inbred lines developed from different tropical-source populations was obtained after eight generation of self-pollination and selection. Among these, 13 inbred lines, originally developed from Malaysian, Indonesian, Hawaiian, Taiwanese and Thailand source-populations, were selected for associate mapping of important agronomic traits. For collection of genomic data, 20 seeds from each inbred line were germinated in jiffy cups and seedlings were grown to the two-leaf stage. Genomic DNA was extracted from young leaves of 20 seedlings per line using the DNeasy® Plant Mini Kit from QIAGEN® following the manufacturer's instruction with minor modification to the washing steps. One hundred and fifty microsatellite regions were chosen from the MaizeGDB (http://www.maizegdb.org.php) based on their polymorphism information contents (PIC) and QTL information reported in previous investigations. Six microsatellite primers out of 150 were not amplified even when tested at different annealing temperatures. The amplifications were performed as described in Kashiani et al [12].

The genetic structure of the inbred lines has been investigated with a model-based analysis. SSR data were used to estimate the LD values among loci and their significance on a genome-wide basis. Linkage disequilibrium (LD) estimates were calculated based on the maximum likelihood estimate (MLE) using an expectation maximization algorithm. Classical (D), standardized (D') and conventional (r2) LD coefficients for any pairs of alleles amplified at two loci on a same chromosome were estimated based on [13]. Population structure consisted of a Q matrix that describes the percent subpopulation parentage was derived using the model-based approach described by Falush et al [14]. Furthermore, the complete set of SSR loci was used to obtain co-ancestry K matrices for all the inbred lines based on the model proposed by Loiselle et al [15-16]. The influence of the genetic diversity structure present in the phenotypic variation of target traits was assessed by means of multiple regression. Association analysis was performed with 10,000 permutation replicates to determine chromosome-wise and genome-wise significance levels of association tests. The significance of marker-phenotype associations was tested using: (i) the fixed general linear model including the population structure coefficients, (ii) the mixed linear model including the Q population structure coefficients and the kinship matric. Marker–trait associations were considered reliable using thresholds of $P \leq 0.05$ and 0.01 marker-wise level and $P \leq 0.05$ experiment-wise level (Bonferroni's correction).

Results and Discussion

Results showed that there was a wide range of agronomic and genetic differences among the inbred lines evaluated. Significant variations in agronomic performance revealed by the inbred lines indicated the presence of high genetic diversity among them. The agronomic performance of these inbred lines has been discussed in details in Kashiani et al [17].

A total of 1347 bands were amplified using the SSR markers, out of which 344 were found to be unique bands among the inbred lines. SSR markers exhibited high Shannon's information index (I), Nei's expected heterozygosity (Nei's), and polymorphic information content (PIC), with mean values of 1.053, 0.586 and 0.582 respectively, indicating their appropriateness in detecting genetic variability among the inbred lines. The genetic relationships among the accessions were investigated using a model-based Bayesian clustering method. For estimation of true number of genetically diverse groups (K) after merging the phenotypic and molecular data, the posterior probability of the data for a given K (1 to 13) was estimated. Results pointed out that the minimum and optimal number of hypotheticals well distinct subgroups present among the inbred lines was equal to five. The ad hoc model used obviously showed that the second order of change in log probability of the data was highest when K was set at five (Figure 1).

In this study, effect of population structure that describes the percent subpopulation parentage was estimated using the model-based approach described by Falush et al [14]. This was due to the fact that the presence of population structure can result in spurious associations, that is, associations between a phenotype and markers that are not linked to any causative loci [18]. Such associations can occur when the agronomic trait varies across subpopulations, thereby increasing the probability that affected individuals are

sampled from particular subpopulations. Any marker allele that is in high frequency in the overrepresented subpopulations will then be associated with the trait [19-20].

Four hundred and forty-six pairs of alleles were found to be in linkage disequilibrium, from which only 50 pairs of loci were separated less than 50cM from each other on the same chromosome. A triangle plot was generated for pairwise LD between marker sites in a hypothetical genome fragment, where pairwise LD values of polymorphic sites are plotted. (Figure 1) displays r2 values and the corresponding p-values from 10000 permutation test. Nineteen significant marker-trait associations were detected with the Q-K-GLM model when significance level was set at $p \leq 0.01$ (Table 1).

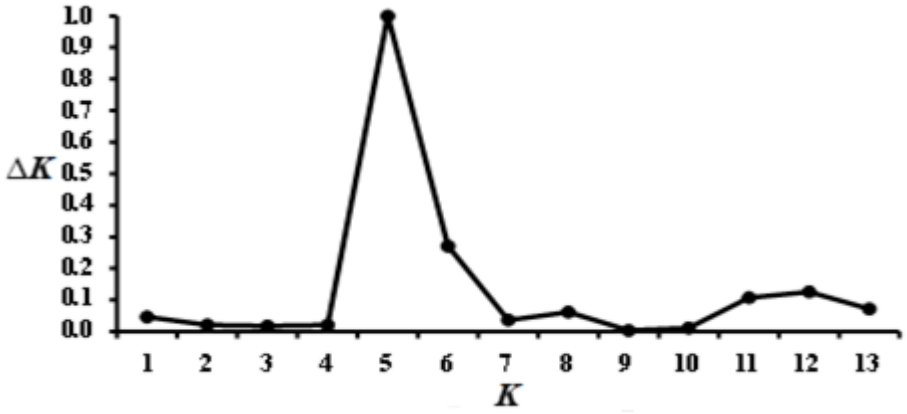

Figure 1: Graph indicating the true number of heterotic groups (K) based on an ad hoc statistic using the normalized second order rate of change in the log probability (ΔK) of data.

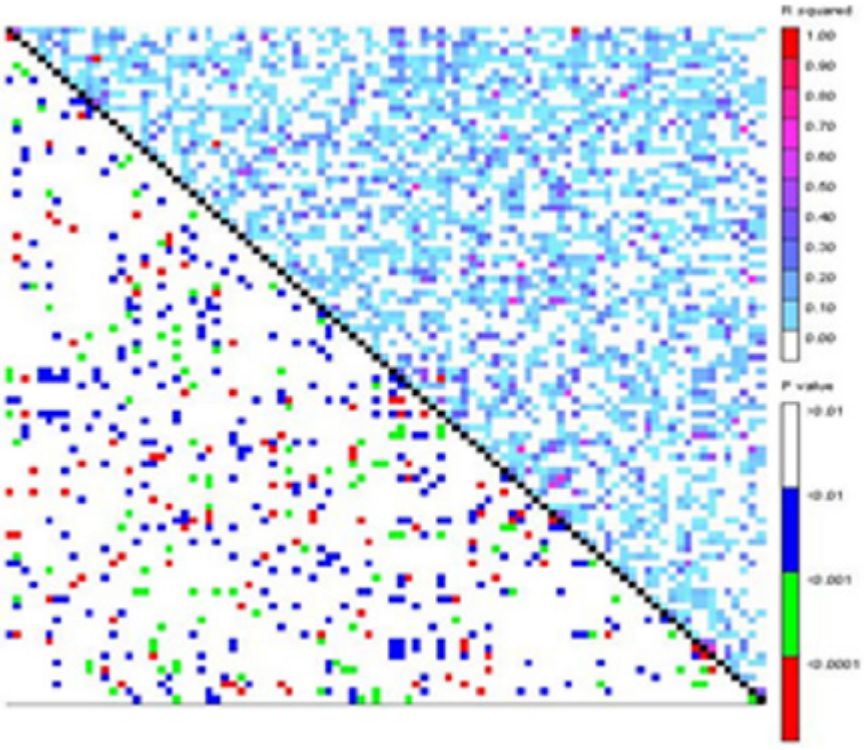

Figure 2: Pairwise LD values of polymorphic sites plotted on both the X- and Y-axis; above the diagonal displays r^2 values and below the diagonal displays the corresponding p-values. Each cell represents the comparison of two pairs of marker sites with the color codes for the presence of significant LD.

Table 1: Nineteen significant marker-trait associations detected with the Q-K-GLM model when significance level was set at p ≤ 0.01.

Trait	SSR	Chr	cM	F Value	P≤F	PP≤F	R^2	MMS	EMS
ANT	umc1279	C9	4.9	50.53	0.0044	0.3596	0.676	3.56	0.07
DHED	bnlg1526	C10	284.7	227.76	0.0005	0.0194*	0.65	56.72	0.25
DHEL	bnlg1518	C10	526.4	2.89E+05	0.0014	0.1012	0.405	3.4	0.01
DHEL	umc1147	C1	714.4	20.72	0.0039	0.2412	0.314	15.8	0.76
DHEL	umc1265	C2	77.7	20.72	0.0039	0.2412	0.314	15.8	0.76
DHEL	umc1652	C4	228.4	17.18	0.0057	0.334	0.353	8.9	0.52
DHEL	bnlg1401	C9	147.5	39.4	0.0063	0.3588	0.397	5	0.13
DHEW	umc1143	C6	17.5	47.13	0.0014	0.0376*	0.512	2073.31	43.99
EH	umc1154	C7	596	66.53	0.0002	0.0182*	0.32	201.88	3.03
EH	umc1426	C7	47.8	72.68	0.0006	0.0240*	0.326	137.13	1.89
EH	umc1153	C5	676.7	38.54	0.0009	0.0456*	0.312	196.71	5.1
EH	bnlg1152	C8	414.1	25.4	0.0024	0.215	0.303	190.7	7.51
HEL	phi053	C3	318.2	56.12	0.001	0.0322*	0.516	11.27	0.2
NE	bnlg1337	C4	708.6	33774.22	0.0042	0.3014	0.548	4.41E+07	1306.15
NE	umc2328	C7	381.5	94.87	0.01	0.574	0.546	5.27E+07	5.56E+05
NKR	bnlg197	C3	544.4	13.64	0.0094	0.42	0.43	86.82	6.37
PH	umc1916	C8	626.7	218.25	0.0005	0.0252*	0.423	667.58	3.06
SIL	umc1019	C5	469.6	14928.07	0.0063	0.3446	0.632	14.94	0.01
TAS	bnlg1798	C3	511.5	30.95	0.009	0.494	0.635	17.77	0.57

Legend: ANT = Anthesis interval, DHED = Dehusked ear diameter, DHEL = Dehusked ear length, DHEL = Dehusked ear weight, EH = Ear height, HEL = Husked ear length, NE = Number of Ears per hectare, NKR = Number of kernels per row, PH = Plant height, SIL = Number of days to silking, TAS = Number of days to tasseling, Chr = Chromosome, cM = Position of SSR based on cM, PP = Permutation probability, MMS = Marker mean squares and EMS = Error mean squares.

However, only seven associations were found to be significant after permutation test. The portion of total variance estimated by the marker ranged from 0.303 to 0.635. Each chromosome possessed at least one (Chromosomes 1, 2 and 6) to three loci (Chromosomes 3 and 7) associated with one agronomic trait. Among the yield related traits, dehusked ear diameter was found to be associated with bnlg1526 located on Chromosome 10, while dehusked ear length was associated with bnlg1518, umc1147, umc1265, umc1652 and bnlg1401 located on Chromosome 10, 1, 2, 4 and 9, respectively. Husked ear length and dehusked ear weight were found to be associated with phi053 (Chr-3) and umc1143 (Chr-6), respectively.

To reduce the percentage of false-positive associations (Type I error), only the associations with significant probability after permutation test were considered significant (pp≤0.05). Based on this criterion, the following associations were considered significant: bnlg1526 associated with dehusked ear diameter, phi053 associated with dehusked ear length, umc1143 associated with dehusked ear weight, umc1154, umc1426 and umc1153 associated with ear height, and umc1916 associated with plant height. Previous studies have shown that sweet corn yield is highly positively correlated with ear characteristics [21-22]. This indicates the usefulness of bnlg1526, phi053 and umc1143 for further molecular breeding for tropical sweet corn.

Conclusion

Based on the results of this study, association mapping revealed the influence of several chromosome regions on the variability of agronomic traits in sweet corn. SSR markers bnlg1526, phi053 and umc1143 were significantly associated with ear characteristics. Therefore, they can be used for marker-assisted selection among the inbred lines toward yield improvement. The markers with significant association with quantitative traits will be further screened on a pool of 40 inbred lines confirmation. Significant associations identified in the present study can also be utilized as a powerful tool for uncovering potential QTL-allele matrix for further marker-assisted breeding on tropical sweet corn.

Acknowledgement

The authors extend their gratitude to Universiti Tunku Abdul Rahman (UTAR) for the financial support provided for this research under UTAR Research Fund (UTARRF), Project No. IPSR/RMC/UTARRF/2022-C2/P01.

Conflict of Interest

None.

References

1. Kashiani P, Saleh G, Osman M, Habibi D (2011) Sweet corn yield response to alternate furrow irrigation methods under different planting densities in a semi-arid climatic condition. African Journal of Agricultural Research 6(4): 1032-1040.
2. Lynch M, Walsh B (1998) Genetics and Analysis of Quantitative Traits. Sunderland, MA: Sinauer Associates.
3. Abu-Sin M, Saleh G, Abdullah NAP, Kashiani P (2020) Genetic diversity among tropical maize inbred lines as revealed by SSR markers. Australian Journal of Crop Science 14(12).
4. Wang Y (2005) Diversity of microsatellite markers in the populations of Picea asperata originating from the mountains of China. Plant Science 168(3): 707-714.
5. Zhang Q, (2008) Characterization of Tomentosa cherry (Prunus tomentosa Thunb.) genotypes using SSR markers and morphological traits. Scientia Horticulturae 118(1): 39-47.
6. Al-Mamun M (2022) Characterization and Genetic Diversity of Photoperiodic among Mutant Kenaf (Hibiscus Cannabinus L.) Using EST-SSR Markers. Journal of Natural Fibers 19(15): 10693-10707.
7. Coetzee R (2004) Characterization of kenaf (Hibiscus cannabinus L.) culfivars in South Africa. Submitted in fulfilment of the requirements of the degree of Magister Scientae Agriculturae.
8. Kashiani P (2012a) Molecular characterization of tropical sweet corn inbred lines using microsatellite markers. Maydica 57: 154-163.
9. Breseghello F, Sorrells ME (2006) Association Analysis as a Strategy for Improvement of Quantitative Traits in Plants. Crop Science 46(3): 1323-1330.
10. Mahender T, Venkatesh B (2018) Optimization of flow forming process parameters of al-8014 using genetic algorithm. International Journal of Engineering & Technology 7(2): 868-873.
11. Nordborg M, Tavare S (2002) Linkage disequilibrium: what history has to tell us. Trends Genet 18(2): 83-90.
12. Kashiani P (2012b) Demarcation of informative chromosomes in tropical sweet corn inbred lines using microsatellite DNA markers. Genetics and Molecular Biology 35: 614-621.
13. Lewontin RC, Kojima K (1960) The evolutionary dynamics of complex polymorphisms. Evolution 14(4): 458-472.
14. Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164(4): 1567-1587.
15. Loiselle BA, Sork VL, Nason J, Graham C (1995) Spatial Genetic Structure of a Tropical Understory Shrub, Psychotria officinalis (Rubiaceae). American Journal of Botany 82(11): 1420-1425.
16. Hardy OJ, Vekemans X (2002) spagedi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. Molecular Ecology Notes 2(4): 618-620.
17. Kashiani P, Saleh G, Abdullah NAP, Sin MA (2014) Evaluation of genetic variation and relationships among tropical sweet corn inbred lines using agronomic traits. Maydica 59(3): 275-282.
18. Lander E, Schork N (1994) Genetic dissection of complex traits. Science 265(5181): 2037-2048.
19. Ewens WJ, Spielman RS (1995) The transmission/disequilibrium test: history, subdivision, and admixture. American Journal of Human Genetics 57(2): 455-464.
20. Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155(2): 945-959.
21. Kashiani P, Saleh G (2010) Estimation of genetic correlations on sweet corn inbred lines using SAS mixed model. American Journal of Agricultural and Biological Sciences 5(3) 309-314.